AWARD NUMBER:
W81XWH-16-1-0176

TITLE:
Chemical Library Screening for Potential Therapeutics Using Novel Cell-Based Models of ALS

PRINCIPAL INVESTIGATOR:
Keith T. Gagnon, Ph.D.

RECIPIENT:
Southern Illinois University
Carbondale, Illinois, 62901

REPORT DATE: June 2017

TYPE OF REPORT:
Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
June 2017	Annual	15 May 2016 - 14 May 2017

4. TITLE AND SUBTITLE

Chemical Library Screening for Potential Therapeutics Using Novel Cell-Based Models of ALS

5a. CONTRACT NUMBER

5b. GRANT NUMBER W81XWH-16-1-0176

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Keith T. Gagnon, Ph.D.

E-Mail: ktgagnon@siu.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Southern Illinois University
1263 Lincoln Dr.
Carbondale, IL 62901

Stanford University
1291 Welch Road
CCSR-0133
Stanford, CA 94305

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The leading cause of inherited amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) is a microsatellite repeat expansion in the C9ORF72 gene. This disorder is referred to as C9FTD/ALS. The disease mechanisms are still poorly understood. However, it is clear that the repeat expansion sequence is made into RNA that can aggregate in the nucleus of patient cells. The expansion RNA can also be translated into repetitive polypeptides in the cytoplasm of patient cells. These two processes are expected to play key roles in the initiation and progression of disease at the molecular and cellular level. Drugs that can block or reverse these processes would hold promise as therapeutics to treat C9FTD/ALS. The overall goal of this project is to develop new cell-based models of C9FTD/ALS that recapitulate these two disease processes. RNA foci and repetitive polypeptides will be visible through fluorescence microscopy. These cells will then be used for high throughput chemical library screening to identify and characterize molecules with therapeutic potential. For this Annual Progress Report, we report on progress for year 1 according to the proposed Statement of Work. The major task of Year 1 was to generate neural stem cells that inducibly express fluorescently-labeled foci and repetitive polypeptides. Toward this task we have successfully engineered inducible expression vectors for the C9FTD/ALS repeat expansion and created neural stem cells that express the tetracycline receptor protein. We are working to overcome unexpected pitfalls and combine these sub-tasks together to complete Milestone 1. We are currently preparing preliminary cell models to send to the Stanford HTBC for initial screening to meet milestones for year 2.

15. SUBJECT TERMS
ALS, repeat expansion, RNA, fluorescence, cell-based models, chemical library, high throughput screening

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
Unclassified	Unclassified	Unclassified	Unclassified	17	19b. TELEPHONE NUMBER *(include area code)*

TABLE OF CONTENTS

1. INTRODUCTION:

We are developing cell-based models of C9FTD/ALS, the most common inherited form of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD), for chemical library screening to identify compounds that have high promise for therapeutic development. The premise of this project is that cell models must recapitulate the two most basic disease processes, formation of repetitive RNA foci and production of repetitive polypeptides, and these processes must be rapidly detectable through microscopy to enable high throughput cell-based screening. In this progress report for year 1, some of the sub-tasks were accomplished according to the Statement of Work, including design and engineering of inducible expression plasmids, cloning of repeat expansion sequences, and selection of neural stem cells stably expressing the tetracycline receptor protein. However, some unexpected technical and personnel challenges were encountered that delayed cell model production. We have worked to solve some of these technical challenges, which resulted in a publication describing optimization of fluorescent RNA aptamers, and have resolved personnel issues. We have also pursued alternative technical aspects of our approach to streamline production and increase our chances of success toward year 1 goals and we have a plan in place to guide success in year 2.

2. KEYWORDS:

Amyotrophic lateral sclerosis, frontotemporal dementia, C9FTD/ALS, repeat expansion, RNA, RAN translation, fluorescence, cell-based models, chemical library, high throughput screening

3. ACCOMPLISHMENTS:

Year 1 was to be defined by the four Subtasks of Aim 1 as well as the first two Subtasks of Aim 2. These Subtasks fell under two Major Tasks and included an expected Milestone for Year 1. To describe accomplishments for year 1, we will name the Major Tasks, Subtasks, and Milestones. We will then describe the major goals each was meant to accomplish and what was actually accomplished. We will then describe training and professional development for students and personnel as well as how findings were disseminated.

Major goals of the project.

The major goal of the project for year 1, as based on the stated SOW, was to generate neural stem cells that express fluorescent RNA foci and fluorescent repeat polypeptides upon treatment with doxycycline. The major goal also included demonstrating that these cell models of C9FTD/ALS exhibit some of the molecular pathology markers expected for C9FTD/ALS. Finally, these cells were to be handed over to the Stanford HTBC before the end of year 1 to begin establishing cell culture conditions and an assay development workflow.

Major Task 1: Establish neuronal cells that inducibly express fluorescent expRNA and poly-dipeptides.

Milestone Expected: Generate C9FTD/ALS cell-based model that recapitulates molecular pathology and is compatible with high throughput screening (*6-9 months, projected completion 01-15-17*). (**60% complete**)

Subtask 1: Build custom expRNA expression vectors (*2-3 months, projected completion 09-01-17*). (**80% complete**)
Subtask 2: Stably transform nH9 cells with TetR (*1 month, projected completion 10-01-17*). (**100% complete**)
Subtask 3: Stably transform nH9-TR cells with custom vectors for expRNA expression (*1 month, projected completion 11-01-17*). (**20% complete**)
Subtask 4: Validate c9FTD/ALS molecular pathology of new cell-based models (*2-4 months, projected completion 02-01-17*). (**20% complete**)

Major Task 2: High throughput screening of chemical libraries at Stanford HTBC.

Milestone Expected: Identification of lead molecules for further validation and therapeutic development (*7-11 months, projected completion 09-15-17*). (**0% complete**)

Subtask 1: Cell culture establishment at HTBC (*1 month, projected completion 03-01-17*). (**0% complete**)

Subtask 2: Assay development and high throughput screen workflow protocol (*1-2 months, projected completion 05-01-17*). (**0% complete**)

Accomplishments under these goals.

Major Task 1, Subtask 1.

Our first major activity in year 1 was the design and engineering of a custom inducible plasmid for expressing fluorescently-labeled repeat expansion RNA and poly-dipeptides. The specific objective was to create a plasmid with features to enhance the likelihood of developing successful cell-based models. These included i) ability to swap fluorescent tags into the plasmid in a modular fashion using traditional restriction enzymes, ii) *in vitro* transcription if needed, iii) a designated site for cloning repeat expansions, and iii) inducible by doxycycline.

Significant results and key outcomes include the accomplishment of most of these objectives and an analysis of the fluorescent Broccoli RNA aptamer that characterized its folding and stability. We designed our expression plasmid based upon a commercially available Tet-On 3G plasmid, called pTRE-3G (Clontech). We modified this plasmid by removing its existing multiple cloning site (MCS) and inserting a custom MCS that was chemically synthesized (**Figure 1A**). This new MCS contained a broader range of compatible restriction endonuclease cleavage sites and a T7 RNA polymerase promoter and terminator sequence for *in vitro* transcription, if needed later (**Figure 1B**). This MCS swapping experiment required site-directed mutagenesis of an existing restriction site elsewhere on the plasmid. We have named the resulting custom plasmid that we have generated pINC3G. To test pINC3G for its ability to support inducible gene expression, we cloned an EGFP protein into the MCS. Upon transient transfection of pINC3G-EGFP into HeLa Tet-On 3G cells (which express the appropriate tetracycline receptor, or TR, protein), we observed strong fluorescence in the presence of doxycycline (**Figure 2A**).

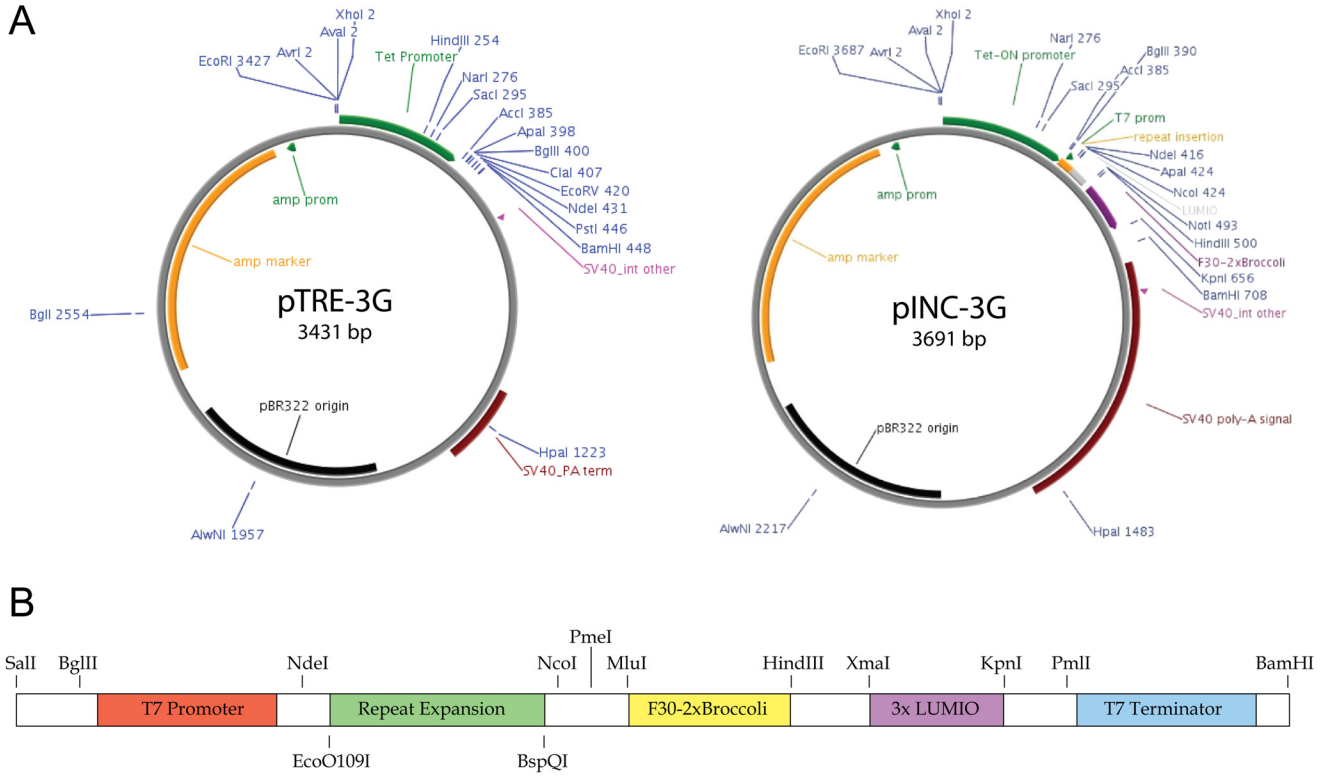

Figure 1. (**A**) Parent plasmid, pTRE-3G, is shown on left and the custom inducible plasmid we have engineered, pINC-3G, is shown on the right. (**B**) Schematic layout of custom MCS used to make pINC-3G and key functional elements.

To prepare the plasmid for fluorescent labeling of RNA foci and poly-dipeptides, which are the hallmarks of disease that will be used as readouts for chemical screening, we inserted a dimeric Broccoli RNA aptamer sequence (Filonov et al., 2014) downstream of the future repeat cloning site (**Figure 1B**). The aptamer tag should enable visualization of repeat expansion RNA foci in live cells. We then inserted a 3x LUMIO tag downstream of the Broccoli tag (**Figure 1B**). The 3x LUMIO tag is in three different reading frames and should

produce a tagged poly-dipeptide no matter what reading frame is used during translation of the repeat expansion RNA (Adams et al., 2002; Irtegun et al., 2011). In order to test the Broccoli aptamer tag, we cloned into pINC3G-dBroc-LUMIO a sequence encoding the non-coding RNAs U24 (snoRNA) or 7sK (snRNA). Upon transfection into cells, however, we saw very poor green fluorescence of the RNA. Expecting that this result is due to improper folding of dimeric Broccoli, we removed dimeric Broccoli from the plasmid and instead inserted two Broccoli aptamers embedded in a stably-folding three-way junction RNA called F30 (Filonov et al., 2015). We have expressed F30-2xBroccoli in HeLa TetOn 3G cells and observed reasonable fluorescence (**Figure 2B**). If F30-2xBroccoli signal still is not bright enough when fused to repeat expansion RNA then we have the option of trying a new aptamer called Mango (Dolgosheina et al., 2014), which has been characterized as potentially more efficient than Broccoli. We will embed two Mango aptamers into the F30 scaffold.

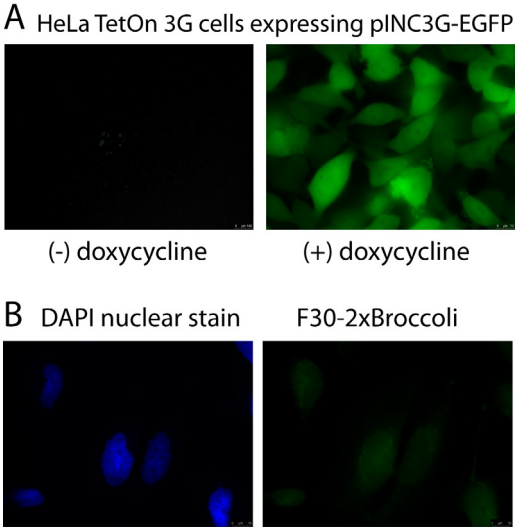

Figure 2. (**A**) Successful testing of pINC-3G plasmid using EGFP expression. (**B**) Successful expression of F30-2xBroccoli pINC-3G plasmid.

Testing of LUMIO tags requires non-canonical translation of repeat expansions in all three frames. Therefore, we began cloning C9FTD/ALS repeat expansions into pINC3G. To do so, we turned to a method known as recursive directional ligation, or RDL (Grabczyk and Usdin, 1999; Meyer and Chilkoti, 2002). In this method, a small 5 repeat sequence flanked by type II restriction enzyme sites is cloned into the plasmid. These repeats were either the sense expansion $(GGGGCC)_5$ or the antisense expansion $(CCCCGG)_5$. The type II restriction sites were chosen based on a previous publication (Mizielinska et al., 2014) and required extensive site-directed mutagenesis of our pINC3G plasmid to remove redundant restriction sites. After cloning, the type II restriction enzymes are used to cut the plasmid and double the number of repeats with each round of RDL. Using this technique we have now generated pINC3G plasmids with 5, 10 or 20 repeats for both sense and antisense expansions (**Figure 3A**). We are currently working on completing the next round of expansion to 40 repeats. The minimum number needed to form foci and generate poly-dipeptide translation products is expected to be between 30-50 repeats (Green et al., 2016; Jain and Vale, 2017). Therefore, the current plasmids we have may not require an additional round of expansion. We will test these plasmids to observe fluorescence of the F30-2xBroccoli-tagged repeat RNA and the LUMIO-tagged poly-dipeptides. While we are hopeful that LUMIO tags will be sufficient, we are prepared to remove the LUMIO tag and insert an mCherry fluorescent protein tag and fuse it in-frame with one of the potential poly-dipeptides.

Figure 3. (**A**) Sequence alignment demonstrating cloning of 5, 10 and 20 repeat expansions into pINC-3G. (**B**) Sequence alignment demonstrating replacement of a canonical ATG start site with a non-canonical ACG start site to support RAN translation.

While preparing our pINC3G plasmids, we realized that our original design was flawed and included a normal canonical AUG start codon. In order for translation to mimic the disease mechanism a canonical start coding should not be present (Green et al., 2016). This information was only recently discovered since the

mechanism at play is only now beginning to be understood. It involves unusual translation called repeat-associated non-AUG (RAN) translation. Therefore, we have now gone back and deleted the canonical translation start codon that originally existed in our plasmid and have redesigned it to contain a non-canonical ACG codon (**Figure 3B**). We believe this will provide the disease-like RAN translation needed to recapitulate disease. All of the plasmid designs and generations we have made have been saved. Therefore, we are able to mix-and-match designs. For example, if a canonical start codon is needed in order to elicit sufficient translation of disease-associated poly-dipeptides, we can easily switch back to that construct.

Other achievements during this first major activity derived from investigating why Broccoli RNA aptamer did not fold properly. Since no other reasonable aptamers were available at the time of project initiation, we felt compelled to try and understand and improve the Broccoli RNA aptamer. We performed a straightforward biochemical investigation of its structure and stability. We performed a systematic mutagenesis screen as well. We uncovered key factors that controlled stability and fluorescence and discovered mutations that made Broccoli aptamer work better, which we named enhanced Broccoli, or eBroccoli. This application-driven project was published during the project period (Ageely et al., 2016) and represents one of the products (**Figure 4A**). We have incorporated these changes into our F30 scaffold and show that the new F30-2x-eBroccoli aptamer also works well (**Figure 4B**). Now that an additional aptamer is available, Mango RNA aptamer, we will consider testing this aptamer as well for its fluorescence when fused to non-coding RNAs or repeat expansion RNA.

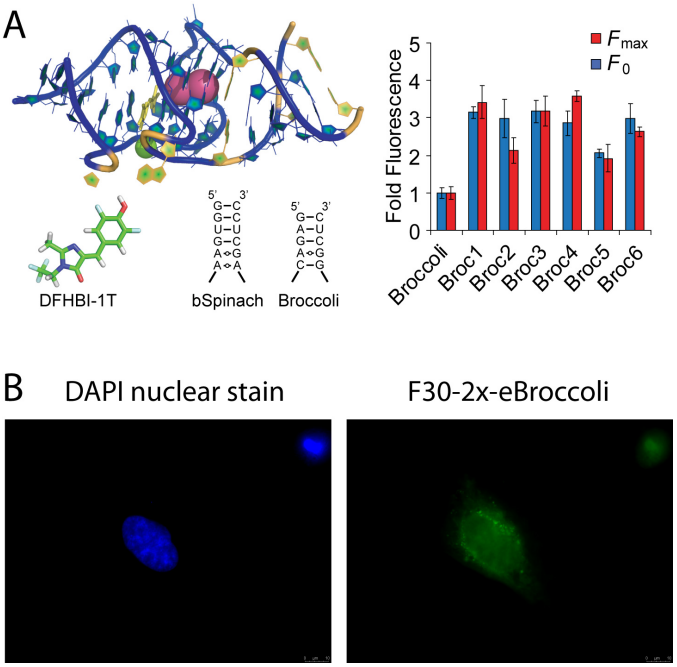

Figure 4. (**A**) Table of content graphic from our published investigation of fluorescent Broccoli RNA aptamer structure and stability. (**B**) Successful expression of F30-2x-eBroccoli pINC-3G plasmid.

Stated goals not met for this Subtask are evidence for working fluorescence aptamers for tagging the repeat expansion RNA and validated fluorescence tags for visualizing the repeat poly-dipeptides. We have also not demonstrated final cloning of repeat expansion sequences. We have a plan to demonstrate these features and also alternative plans in case they fail. These are outlined below in the CHANGES/PROBLEMS section.

Major Task 1, Subtask 2.

Our second major activity in year 1 was to generate human neural stem cells that stably express a tetracycline receptor (TetR, or TR) protein. The specific objective was to create cells that would support doxycycline-inducible expression of the pINC3G plasmids we engineered. These cells are not commercially available and so had to be custom made.

Significant results and key outcomes were the establishment of neural stem cell culture and demonstration of stable expression of the TR gene in these cells without significantly impact key gene expression markers of the neural cell lineage, proliferation and multipotency. We obtained pCMV-Tet3G (Clontech) plasmid containing the appropriate TR gene for Tet-ON 3G regulation. We obtained neural stem cells derived from human embryonic H9 stem cells. We have named these cells nH9 cells. We transfected the pCMV-Tet3G plasmid into nH9 cells and then performed antibiotic selection with neomycin. Surviving colonies were selected, expanded and tested for their ability to i) express sufficient levels of the TR gene and ii) maintain gene expression markers indicative of the parental cell lines. Upon isolation and qPCR we identified a few clonal populations that grow similarly to the parental nH9 cells and expressed the TR gene at sufficient levels (**Figure 5A**). These cells are now prepared and ready for the next stage of cell model development. There are no other achievements to discuss for this Subtask. There are no stated goals not met for this Subtask.

Major Task 1, Subtask 3.

Our third major activity in year 1 was to generate nH9-TR cells (stably expressing the TR gene) that also contained stably-transfected pINC3G plasmids. The specific objective is to avoid the need to repeatedly transfect cells with the repeat expansion expressing plasmids, which is not feasible for high throughput

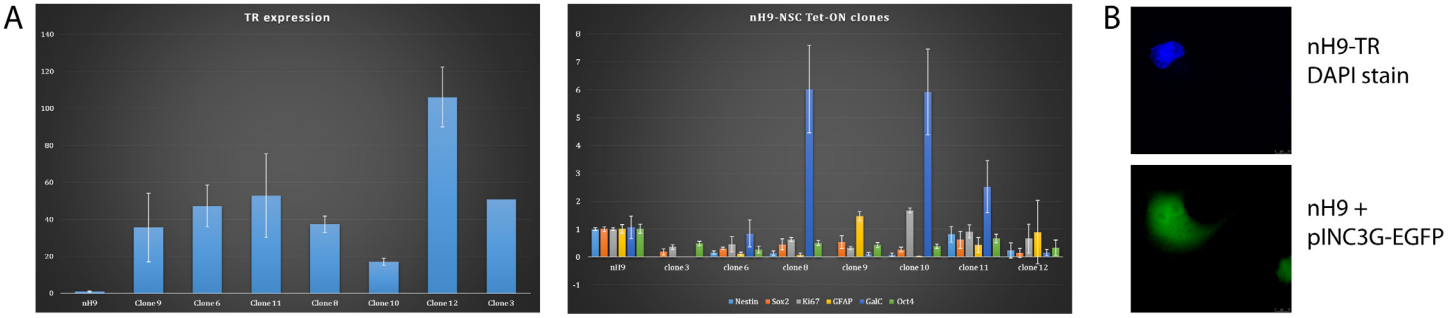

Figure 5. (**A**) Expression of the TR gene in nH9-TR clonal cell lines (left) and expression levels for genes that are markers for neuronal cell state, multipotent stem cell state, and proliferation state. Clone 11 was the most promising (**B**) Fluorescence microscopy demonstrating successful induction of expression from pINC-3G-EGFP in nH9-TR(11) cells.

chemical library screening. These cells, therefore, would contain all of the necessary disease-associated genetic information, but would not present any of the disease-associated biomarkers, including foci and poly-dipeptides, until these are induced with doxycycline.

Significant results and key outcomes are the successful testing of pINC3G-EGFP by transient transfection in nH9-TR cells (**Figure 5B**). These experiments demonstrate that our nH9-TR cells are compatible with our pINC3G plasmids. However, stated goals not met for this Subtask are the generation of nH9-TR cells stably and inducibly expressing fluorescently-tagged repeat RNA and poly-dipeptides. Thus, once we have completed and validated custom plasmids, we will perform stable cell line selection. We have the option of using hygromycin or puromycin for our selections. We expect that stable cell line selection will be straightforward and only require the 1 month stated in our SOW. There are no other achievements to discuss for this Subtask.

Major Task 1, Subtask 4.

Our fourth major activity in year 1 was to validate the cell-based models prepared in Subtasks 1-3 of Aim 1. The specific objective for this activity is to ensure that cell-based models are faithful representations of cellular disease. Positive results for other disease markers, including rRNA defects or TDP-43 mislocalization to the cytoplasm, would be desirable. However, these markers are not required and will not prevent successful screening with chemical libraries as long as fluorescent disease biomarkers established in Subtasks 1-3 are successful.

Significant results and key outcomes are the testing of repeat expansion RNA foci detection with fluorescence in situ hybridization (FISH) on existing cell-based models, a key assay in helping to validate our cell lines. We used patient-derived fibroblast cells and performed a FISH protocol to detect both sense and antisense foci (**Figure 6**). For our cell-based nH9-TR model-cells, we expect to see foci in live cells if our fluorescent aptamer tagging is successful. However, FISH is a gold standard that can validate these foci or detect them in case aptamers continue to be problematic. FISH is also a technique that can be automated with robotics at the Stanford HTBC if necessary. Therefore, it is important to establish this protocol for disease marker detection. Stated goals not met for this Subtask are the validation of cell-based models. There are no other achievements to discuss for this Subtask.

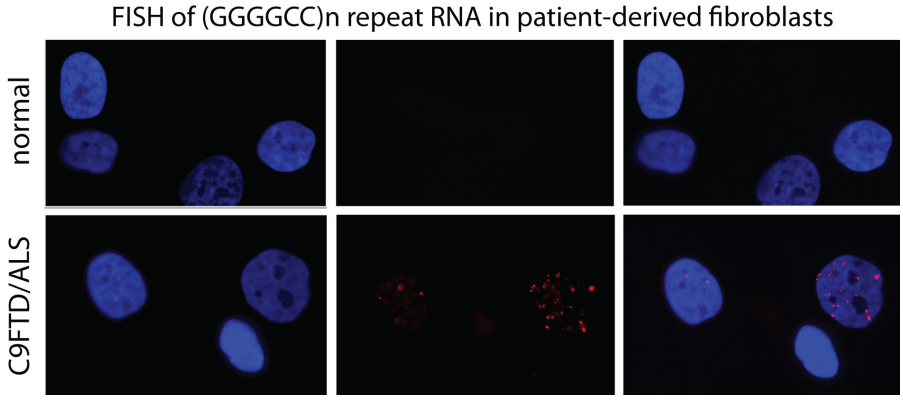

Figure 6. FISH to detect repeat expansion RNA foci (red) associated with C9FTD/ALS. Probing in normal fibroblast cells (top panels) and C9FTD/ALS patient-derived fibroblasts (bottom panels). Blue, DAPI stain.

Major Task 2, Subtask 1 and 2.

Our fifth major activity planned for year 1 was establishing cell culture and assay development at the Stanford High Throughput Bioscience Center (HTBC) using our cell-based models. The specific objective of this activity is to prepare practical protocols for high throughput chemical library screening. In our experience to

this point nH9 cells are quite straightforward to grow and maintain. The only additional step of note is that cell culture dishes must be coated with a biological matrix, such as Geltrex or fibronectin, prior to seeding cells. There are no significant results to describe for this activity. The stated goals for this activity have not been met yet and there are no other achievements to discuss for these Subtasks.

Opportunities for training and professional development.

A variety of students and young professionals have participated in this project to date. They include an undergraduate student, four graduate students, and a postdoctoral fellow. The postdoctoral fellow, Dr. Maria Barton, started working on this project in February, 2017 (4 months ago) and has learned several new techniques, including stem cell culture and FISH. We have a new PhD student joining the laboratory this summer, Ms. Katy Ovington, who will be working with Dr. Barton. This will provide an opportunity for Dr. Barton to teach skills to Katy and improve her teaching abilities.

An undergraduate student, Mr. James Donohue, assisted one of the graduate students this year, Ms. Kushal Rohilla, and learned how to clone repeat expansions. This training under the graduate student provided a fantastic opportunity for Kushal to learn how to teach molecular biology techniques and James learned important research skills that will serve him well in his future career. The other three graduate students who worked full-time or part-time on this project learned new skills in cell culture and molecular cloning.

Ms. Kushal Rohilla attended an international conference, RNA Metabolism in Neurological Disorders, in San Diego, California and was able to present our progress on this project. She also presented her research at a local SIU School of Medicine symposium and won a first place poster prize. Another one of the graduate students, Mr. Zachary Kartje, attended a conference, The Oligonucleotide Therapeutics Society Meeting, in Montreal, Canada and was able to present our progress on this project. Zachary also won a poster prize at this conference. Mr. Christopher Barkau is a graduate student that only worked part-time on the project and Mrs. Ayomi Hewavidana, a new graduate student, learned several important new assays, including cell culture and FISH.

Dissemination of results to communities of interest.

Results were disseminated by conference presentations, abstracts, and publications. Two conferences in 2016 and one in 2017 were used to disseminate research. An oral presentation was given at one conference and poster presentations given at the other two meetings. One publication in 2016 and a review article in 2017 have been used to disseminate results. Conference and publication abstracts are provided in Appendix I.

Plans to accomplish goals during the next reporting period.

The next reporting period has already began. Our priority is to complete plasmid cloning so that we can finalize our cell based models. However, to get back on schedule we plan to immediately send nH9 cells to the Stanford HTBC so that they can begin assay development (Subtask 1 of Major Task 2). Approximately one month will be needed to establish cell culture conditions that are compatible with high throughput screening. During this period we will continue to finish Subtasks that are behind schedule, specifically Subtasks 1, 3 and 4 of Major Task 1.

To ensure no further delays, we are prepared to make changes to the future experimental goals. For example, if LUMIO tags do not yield favorable fluorescence we will immediately swap in an mCherry fluorescent protein tag in frame with one of the putative RAN translation reading frames. This approach has been used by others with EGFP fluorescent protein. Unlike LUMIO, it will not allow us to visualize all potential reading frames, but the compromise will be fair since it will guarantee strong fluorescence for at least one poly-dipeptide. The poly-dipeptide that is currently in frame with the near cognate start codon (ACG) is poly-glycyl-alanine (poly-GA) and that is the reading frame we will use for mCherry.

We are also prepared to abandon fluorescent RNA aptamers if they continue to be inefficient. The Stanford HTBC is capable of performing high throughput screening using FISH protocols. We will turn to this workflow if needed since we have a reliable FISH protocol in hand. As part of Subtask 1 of Major Task 2 we will ask the HTBC to also establish FISH conditions in anticipation that fluorescent aptamers may not be sufficient.

Our primary goal for the next reporting period is to get suitable cells into HTBC screening without compromising the overall objective of the project. As we finish preparing cell-based models this summer we will work with the HTBC to move forward.

4. IMPACT:

Impact on the development of the principal disciplines of the project.

Our interim progress on cell-based models was presented as a poster at the RNA Metabolism in Neurological Disease meeting in November of last year. The audience was interested in our approach and we have had several inquiries regarding accessibility of the models when we are done, as well as interest in collaborations. When we publish our cell-based models and the subsequent chemical library screening results, we believe we will make an important impact by shifting research practices. Currently there are no easy-to-access or easy-to-use cell-based models that are representative of the disease. We believe our work will fill this gap. There are also no informative studies using chemical libraries to identify lead compounds for C9FTD/ALS, so we believe we will add a tremendous contribution to the principal discipline when this project is complete by providing detailed insight into potential therapeutics.

We believe our innovative approach of dual fluorescence to image both RNA foci and poly-dipeptides simultaneously in live neuronal cells will transform the standard for basic research in the field. Our characterization of the many moving parts involved in engineering successful cell-based models will provide a foundation for researchers in the principal discipline to build upon. Our cell-based models will also make it significantly easier for researchers outside of the field to enter in and make important contributions.

Impact on other disciplines.

We have published an aspect of this project regarding optimization of fluorescent RNA aptamers, which came as a necessity to improve imaging capabilities. That work was presented recently at a conference as a poster and was well-accepted by the research community, indicating that we are making an impact. It was appreciated by many that the original Broccoli RNA aptamer was inefficient and that our work was important to improve its application. The published manuscript has already been cited by others in the field of fluorescent RNA aptamers.

Impact on technology transfer.

There is nothing to report for this section.

What was the impact on society beyond science and technology?

There is nothing to report for this section.

5. CHANGES/PROBLEMS:

Changes in approach and reasons for change.

No significant changes to our approach have been taken, other than alternative approaches described in the original proposal. We have had to reconsider the use of specific fluorescent RNA aptamers and fluorescent protein tags, which was anticipated. We have had to redesign the custom repeat expansion expression plasmid due to some unexpected technical challenges (described above in ACCOMPLISHMENTS). We have had to adjust our timeline for establishing cell culture assays at the HTBC. However, none of these changes have led to significant departure from our overall approach or goals.

Actual or anticipated problems or delays, and actions or plans to resolve them.

We have encountered a few problems during the course of this reporting year that have resulted in delays that were not anticipated. These include technical challenges and personnel issues.

Technical issues arose primarily regarding the engineering of custom plasmids. To insert our custom MCS into the pTRE3G plasmid we had to perform site-directed mutagenesis to change a redundant restriction enzyme site. We had to perform additional rounds of mutagenesis to also delete or alter several cryptic restriction sites that (unexpectedly) matched our type II restriction enzymes used for repeat expansion cloning. Site directed mutagenesis takes at least a week for each round, thus setting us back several weeks. We found that the Broccoli aptamer, as published, was inadequate in our human cells. We modified Broccoli by placing in

another scaffold, the F30 3-way junction motif, and have hopefully solved this problem. We have also realized that our repeat expansion expression plasmid may not be suitable for stable cell line selection because the repetitive repeat sequence may result in epigenetic silencing of the transgene. Therefore, we are circumventing this problem by placing the core elements of our pINC3G plasmid into the backbone of a transposase-mediated recombination plasmid, called piggyBac.

For making our cell-based model, we expected that live-cell dyes that can highlight the cell nucleus would be compatible, such as Hoechst dye. However, we have discovered that these dyes are somewhat toxic to the nH9 neural cells and do not allow them to continue to divide. We therefore have set out to solve this problem by selecting neurons that constitutively express a blue fluorescent protein that localizes to the nucleus. This protein is called TagBFP and it has a nuclear localization signal on it. We have tested this plasmid by transient transfection to confirm that it works (**Figure 7**). We believe this approach will solve our problem and enable long-term live-cell imaging.

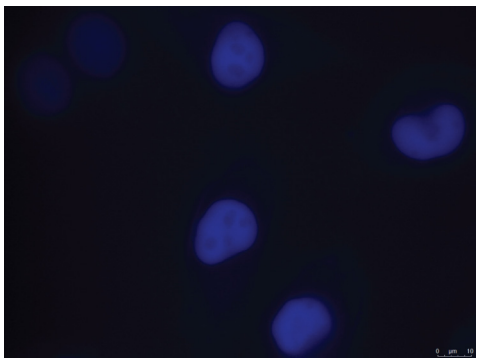

Figure 7. Transient tag-BFP-NLS protein expression in HeLa TetOn 3G cells showing blue fluorescence in the nucleus. This fluorescent protein can replace the need for toxic nuclear dyes like Hoechst.

We also had an unexpected cell culture contamination of our nH9 cells that persisted for 6 weeks. This ultimately was due to a contaminated water source in the laboratory. We identified the bacterium as *Ralstonia insidiosa*, a particularly difficult soil microbe to get rid of due to multiple antibiotic resistance and spores that resist high temperatures and high concentrations of ethanol. We have addressed this issue and feel confident that it will not recur.

Personnel issues arose regarding the hiring of a postdoctoral fellow and a graduate student to work on this project. Postdoc hiring was scheduled for mid-November 2016. However, due to visa issues Dr. Maria Barton was unable to start working until mid-February. This resulted in a 3 month delay of hire for a highly skilled and critical member of the research team. In the fall of 2016 a new graduate student, Mrs. Ayomi Hewavidana, joined the laboratory and was assigned to assist with this project. Due to personal issues and health concerns she was unable to perform her laboratory duties. Because of the nature of these problems it was not possible to address them directly or early. This resulted in an unproductive 7 months and this student has subsequently left the graduate program.

We do not expect special treatment for our technical and personnel issues and understand that the proposed work must be completed. Nonetheless, these very real issues plagued our daily work on this project and are the reason why we have fallen behind our proposed SOW schedule. We hope to make up lost ground in the coming months by focusing our efforts and getting preliminary cell models to the HTBC so they can begin establishing robotics protocols.

Changes that had a significant impact on expenditures.

The technical and personnel issues we encountered during year 1 had a significant impact on expenditures. They resulted in less funds being spent. This is due to the delayed hiring of a postdoctoral fellow and a delay in sending cells to the HTBC for preliminary assay development. We anticipate that these funds will be spent as we make a push to begin assay development and setup cell culture protocols at the Stanford HTBC this summer.

Significant changes in use or care of human subjects

There is nothing to report for this section. Human subjects are not applicable to this proposal.

Significant changes in use or care of vertebrate animals.

There is nothing to report for this section. Vertebrate animals are not applicable to this proposal.

Significant changes in use of biohazards and/or select agents

There is nothing to report for this section.

6. PRODUCTS:

Journal publications.

Ageely, E.A., Kartje, Z.J., Rohilla, K., Barkau, C.L., and Gagnon, K.T. (2016) Quadruplex-flanking stem structures modulate the stability and metal ion preferences of RNA mimics of GFP. *ACS Chem. Biol.*, 11:2398-2406. Support acknowledged. http://pubs.acs.org/doi/abs/10.1021/acschembio.6b00047

Rohilla, K.J., and Gagnon, K.T. (2017) RNA Biology of Disease-Associated Microsatellite Repeat Expansions. *in review*. Support acknowledged.

Books or other non-periodical, one-time publications.

There is nothing to report for this section.

Other publications, conference papers, and presentations.

4th RNA Metabolism in Neurological Diseases Conference, November 10-11, 2016, San Diego, CA. Poster presentation. "Cell-based models of repeat expansion disease for cellular and molecular biochemistry."

Invited Seminar, Department of Chemistry, St. Louis University, St. Louis, Missouri, February, 2017. Oral presentation. "Broccoli and CRISPR and What's Cooking in the RNA Kitchen."

27th Annual Trainee Research Symposium, Southern Illinois University School of Medicine, Carbondale, IL, April 21, 2017. Poster presentation. "Cell-Based Models of Repeat Expansion Disease for Cellular and Molecular Biochemistry."

22nd Annual Meeting of the RNA Society, May 30 - June 4th, Prague, Czech Republic. Poster presentation. "Structure-Function Studies of Broccoli RNA Aptamer for Better Fluorescence."

Website(s) or other Internet site(s).

There is nothing to report for this section.

Technologies or techniques

There is nothing to report for this section.

Inventions, patent applications, and/or licenses

There is nothing to report for this section.

Other Products

There is nothing to report for this section.

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

Individuals that have worked on the project.

Name:	Keith T. Gagnon
Project Role:	PI
Researcher Identifier (e.g. ORCID ID):	N/A
Nearest person month worked:	1
Contribution to Project:	Dr. Gagnon supervised and oversaw research progress, experimental troubleshooting, and dissemination of findings.
Funding Support:	Southern Illinois University School of Medicine, DoD ALSRP (this award)

Name:	Kushal J. Rohilla
Project Role:	Graduate Student
Researcher Identifier (e.g. ORCID ID):	N/A
Nearest person month worked:	3
Contribution to Project:	Ms. Rohilla has performed design and cloning of inducible repeat expansion expression plasmids.
Funding Support:	Judith and Jean Pape Adams ALS Research Grant, DoD ALSRP (this award)

Name:	Christopher L. Barkau
Project Role:	Graduate Student
Researcher Identifier (e.g. ORCID ID):	N/A
Nearest person month worked:	3
Contribution to Project:	Mr. Barkau has performed cloning of inducible repeat expansion expression plasmids.
Funding Support:	SIU Graduate Research Fellowship, DoD ALSRP (this award)

Name:	Zachary J. Kartje
Project Role:	Graduate Student
Researcher Identifier (e.g. ORCID ID):	N/A
Nearest person month worked:	9
Contribution to Project:	Mr. Kartje has performed cloning of inducible repeat expansion expression plasmids, stem cell culture, and selection of nH9-TR stable cells.
Funding Support:	SIU Teaching Assistantship, DoD ALSRP (this award)

Name:	Ayomia Hewavidana
Project Role:	Graduate Student
Researcher Identifier (e.g. ORCID ID):	N/A
Nearest person month worked:	9
Contribution to Project:	Mrs. Hewavidana performed cloning of expression plasmids, cell culture, and FISH.
Funding Support:	DoD ALSRP (this award)

Name:	Maria Barton
Project Role:	Postdoctoral Fellow
Researcher Identifier (e.g. ORCID ID):	N/A
Nearest person month worked:	3
Contribution to Project:	Dr. Barton has performed stem cell culture, cloning of repeat expansion expression plasmids, and FISH.
Funding Support:	DoD ALSRP (this award)

Changes in the active other support of the PI since the last reporting period.

Previously funded grants that have completed:

School of Medicine Team Development Grant (TDG)

Gagnon (PI) 0.15 CY person-months $15,000 01/01/15 - 12/31/15
Title: "Demystifying Aggregation in c9FTD/ALS to Enable Therapeutic Development"

ALS Association Starter Grant
Gagnon (PI) 0 CY person-months $40,000 09/01/15 - 08/31/16
Title: "Flexible and Accessible Cell-Based Models of c9FTD/ALS"

Newly funded grants that have started:

Judith and Jean Pape Adams Foundation ALS Research Grant
Gagnon (PI) 0 CY person-months $60,000 02/01/17 - 01/31/18
Title: "C9ORF72 transcription and splicing as therapeutic targets for a genetic form of ALS"

Other partnering organizations.

We have partnered with Stanford University's High Throughput Bioscience Center (HTBC) to provide high throughput chemical library screening as a service. They will provide us this service as part of the proposed research. The contact person at HTBC is the facility director Dr. David Solow-Cordero. The HTBC did not provide any services in this year 1 reporting period. However, they will be providing their service in the next reporting phase.

Organization Name:	Stanford University High Throughput Bioscience Center
Organization Location:	Stanford, California
Partners Contribution:	
Financial Support:	None.
In-kind Support:	None.
Facilities:	Provides high throughput robotics facility as a service for the chemical library screening phase of this project.
Collaboration:	The facility staff will help the project staff rank results and interpret results to identify promising lead compounds.
Personnel Exchanges:	One project staff from the PI institution may travel to the HTBC facility in reporting year 2 to help establish cell culture and assay conditions.

8. SPECIAL REPORTING REQUIREMENTS

COLLABORATIVE AWARDS:

Nothing to report.

QUAD CHARTS:

Nothing to report.

9. APPENDICES:

See Appendix I below for publication and meeting abstracts pertaining to the year 1 reporting period. See Appendix II below for additional references pertaining to the year 1 report.

APPENDIX I.

Ageely, E.A., Kartje, Z.J., Rohilla, K., Barkau, C.L., and Gagnon, K.T. (2016) Quadruplex-flanking stem structures modulate the stability and metal ion preferences of RNA mimics of GFP. *ACS Chem. Biol.*, 11:2398-2406. Support acknowledged. http://pubs.acs.org/doi/abs/10.1021/acschembio.6b00047

Abstract:
The spinach family of RNA aptamers are RNA mimics of green fluorescent protein (GFP) that have previously been designed to address the challenges of imaging RNA inside living cells. However, relatively low levels of

free intracellular magnesium limited the practical use of these aptamers. Recent cell-based selections identified the broccoli RNA aptamer, which requires less magnesium for fluorescence, but the basis for magnesium preference remained unclear. Here, we find that the broccoli RNA structure is very similar to that of baby spinach, a truncated version of the spinach aptamer. Differences in stability and metal ion preferences between these two aptamers, and among broccoli mutants, are primarily associated with the sequence and structure of predicted quadruplex-flanking stem structures. Mutation of purine−purine pairs in broccoli at the terminal stem−quadruplex transition caused reversion of broccoli to a higher magnesium dependence. Unique duplex-to-quadruplex transitions in GFP-mimic RNAs likely explain their sensitivity to magnesium for stability and fluorescence. Thus, optimizations designed to improve aptamers should take into consideration the role of metal ions in stabilizing the transitions and interactions between independently folding RNA structural motifs.

Rohilla, K.J., and Gagnon, K.T. (2017) RNA Biology of Disease-Associated Microsatellite Repeat Expansions. *in review*. Support acknowledged.

Abstract:
Microsatellites, or simple tandem repeat sequences, occur naturally in the human genome and may have important roles in genome evolution and function. However, the expansion of microsatellites is associated with over two dozen neurological diseases. A common denominator for most of these disorders is the expression of expanded tandem repeat-containing RNA, referred to as "xtrRNA" in this review, which can contribute to molecular pathology in multiple ways. Repeat expansion disorders that express xtrRNA include myotonic dystrophy type 1 (DM1) and type 2 (DM2), several spinocerebellar ataxia (SCA) disorders, Huntington's disease (HD), Fragile X-associated tremor/ataxia syndrome (FXTAS), spinal and bulbar muscular atrophy (SBMA), or Kennedy's disease, and *C9ORF72*-associated frontotemporal dementia (FTD) and amyotrophic lateral sclerosis (ALS), commonly called C9FTD/ALS. This review broadly discusses the impact that simple tandem repeats can have on the biology and metabolism of RNA that contain these aberrant repeat expansions and underscores important gaps in understanding. Merging the molecular biology of repeat expansion disorders with the current understanding of splicing, transcription, transport, turnover and translation of RNA will improve understanding of molecular disease pathology and the identification of new therapeutic options.

4th RNA Metabolism in Neurological Diseases Conference, November 10-11, 2016, San Diego, CA. Poster presentation. "Cell-based models of repeat expansion disease for cellular and molecular biochemistry."

Abstract:
Access to simple cell-based models of neurological repeat expansion disease is critical for investigating biochemical mechanisms and for early therapeutic discovery. Most cell-based models for neurological disease, in particular repeat expansion diseases, are hard to access, can be challenging to use, or else do not sufficiently recapitulate disease at the cellular level. To help bridge this gap, we are engineering straightforward cell-based models of c9FTD/ALS, the leading genetic cause of frontotemporal dementia (FTD) and amyotrophic lateral sclerosis (ALS), designed to allow inducible expression of repeat expansions that can be tracked at the RNA and protein level by fluorescent and affinity tags. This model system should be readily amenable to other repeat expansion disorders. Our cell-based models incorporate the tetracycline receptor gene into commercially available human neural stem cells. These neural stem cells can then be transiently or stably transfected with vectors expressing repeat expansion sequences behind a tetracycline inducible promoter. We have built custom inducible plasmids and are establishing reliable protocols for repeat expansion cloning and expression. Repeat RNA and poly-dipeptides can be fused to modular tags, including small Broccoli RNA aptamers and tetra-cysteine peptide tags for fluorescence or affinity tags for purification. Expression levels are then controlled by doxycycline, which should allow studies like temporal expression and localization as well as step-by-step characterization of disease mechanism at the cellular and biochemical level.

27th Annual Trainee Research Symposium, Southern Illinois University School of Medicine, Carbondale, IL, April 21, 2017. Poster presentation. "Cell-Based Models of Repeat Expansion Disease for Cellular and Molecular Biochemistry."

Abstract:
Access to simple cell-based models of neurological repeat expansion disease is critical for investigating biochemical mechanisms and for early therapeutic discovery. Most cell-based models for these diseases are hard to access, can be challenging to use, or else do not sufficiently recapitulate disease at the cellular level. To help bridge this gap, we are engineering straightforward cell-based models of c9FTD/ALS, designed to

allow inducible expression of repeat expansions that can be tracked at the RNA and protein level by fluorescent and affinity tags. This model system should be readily amenable to other repeat expansion disorders. Our cell-based model incorporates the tetracycline receptor gene into commercially available and relatively easy-to-use human neural stem cells. These neural stem cells can then be transiently or stably transfected with vectors expressing repeat expansion sequences behind a tetracycline inducible promoter. We have built custom inducible plasmids and are establishing reliable protocols for repeat expansion cloning and expression. Repeat RNA and poly-dipeptides can be fused to modular tags, including small Broccoli RNA aptamers and tetra-cysteine peptide tags for fluorescence or affinity tags for purification. Expression levels are then controlled by doxycycline, which should allow temporal studies of expression and localization as well as step-by-step characterization of disease mechanism at the cellular and biochemical level.

22nd Annual Meeting of the RNA Society, May 30 - June 4th, Prague, Czech Republic. Poster presentation. "Structure-Function Studies of Broccoli RNA Aptamer for Better Fluorescence."

Abstract:

The Spinach family of aptamers are RNA mimics of green fluorescent protein (GFP) that have previously been designed to address the challenges of imaging RNA inside living cells. However, relatively low levels of free intracellular magnesium limit the practical use of these aptamers. New cell-based selections identified the Broccoli RNA aptamer, which required less magnesium for fluorescence. However, the structure of Broccoli and the basis for lower magnesium dependence were unknown. Here we find that Broccoli RNA shares the same core quadruplex structure as Spinach and is nearly identical to the structure of a truncated version, Baby Spinach. Differences in stability and metal ion preferences between these two aptamers, and among Broccoli mutants tested, are primarily associated with the sequence, structure and stability of predicted quadruplex-flanking stem and stem-loop structures. Mutation of purine-purine pairs in Broccoli at the terminal stem-to-quadruplex transition caused reversion of Broccoli to a higher magnesium dependence. Unique duplex-to-quadruplex transitions in GFP-mimic RNAs likely explain their sensitivity to magnesium and certain other metal ions. Thus, optimizations designed to improve aptamers should pay careful attention to the role of transitions between distinct or independently folding RNA structural motifs. Systematic mutagenesis and comparative structure-function analyses have allowed us to rationally design an enhanced Broccoli aptamer, called eBroccoli, that exhibits better folding and higher stability than the original Broccoli aptamer.

APPENDIX II.

References

Adams, S.R., Campbell, R.E., Gross, L.A., Martin, B.R., Walkup, G.K., Yao, Y., Llopis, J., and Tsien, R.Y. (2002). New biarsenical ligands and tetracysteine motifs for protein labeling in vitro and in vivo: synthesis and biological applications. J Am Chem Soc *124*, 6063-6076.

Ageely, E.A., Kartje, Z.J., Rohilla, K.J., Barkau, C.L., and Gagnon, K.T. (2016). Quadruplex-Flanking Stem Structures Modulate the Stability and Metal Ion Preferences of RNA Mimics of GFP. Acs Chem Biol *11*, 2398-2406.

Dolgosheina, E.V., Jeng, S.C., Panchapakesan, S.S., Cojocaru, R., Chen, P.S., Wilson, P.D., Hawkins, N., Wiggins, P.A., and Unrau, P.J. (2014). RNA mango aptamer-fluorophore: a bright, high-affinity complex for RNA labeling and tracking. Acs Chem Biol *9*, 2412-2420.

Filonov, G.S., Kam, C.W., Song, W., and Jaffrey, S.R. (2015). In-gel imaging of RNA processing using broccoli reveals optimal aptamer expression strategies. Chem Biol *22*, 649-660.

Filonov, G.S., Moon, J.D., Svensen, N., and Jaffrey, S.R. (2014). Broccoli: rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution. J Am Chem Soc *136*, 16299-16308.

Grabczyk, E., and Usdin, K. (1999). Generation of microgram quantities of trinucleotide repeat tracts of defined length, interspersion pattern, and orientation. Anal Biochem *267*, 241-243.

Green, K.M., Linsalata, A.E., and Todd, P.K. (2016). RAN translation-What makes it run? Brain Res.

Irtegun, S., Ramdzan, Y.M., Mulhern, T.D., and Hatters, D.M. (2011). ReAsH/FlAsH labeling and image analysis of tetracysteine sensor proteins in cells. Journal of visualized experiments : JoVE.

Jain, A., and Vale, R.D. (2017). RNA phase transitions in repeat expansion disorders. Nature *546*, 243-247.

Meyer, D.E., and Chilkoti, A. (2002). Genetically encoded synthesis of protein-based polymers with precisely specified molecular weight and sequence by recursive directional ligation: examples from the elastin-like polypeptide system. Biomacromolecules *3*, 357-367.

Mizielinska, S., Gronke, S., Niccoli, T., Ridler, C.E., Clayton, E.L., Devoy, A., Moens, T., Norona, F.E., Woollacott, I.O., Pietrzyk, J.*, et al.* (2014). C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins. Science *345*, 1192-1194.